

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Bandman, Olga
Corley, Neil C.
Guegler, Karl J.
- (ii) TITLE OF THE INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 - (B) STREET: 3174 Porter Drive
 - (C) CITY: Palo Alto
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94304
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: To Be Assigned
 - (B) FILING DATE: Herewith
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Billings, Lucy J.
 - (B) REGISTRATION NUMBER: 36,749
 - (C) REFERENCE/DOCKET NUMBER: PF-0333 US
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 415-855-0555
 - (B) TELEFAX: 415-845-4166
 - (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 448 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CORNNOT01
 - (B) CLONE: 45517
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	Pro	Gly	Ile	Lys	Arg	Ile	Leu	Thr	Val	Thr	Ile	Leu	Ala	Leu	Cys
1				5					10					15	
Leu	Pro	Ser	Pro	Gly	Asn	Ala	Gln	Ala	Gln	Cys	Thr	Asn	Gly	Phe	Asp
			20					25					30		
Leu	Asp	Arg	Gln	Ser	Gly	Gln	Cys	Leu	Asp	Ile	Asp	Glu	Cys	Arg	Thr
		35					40					45			
Ile	Pro	Glu	Ala	Cys	Arg	Gly	Asp	Met	Met	Cys	Val	Asn	Gln	Asn	Gly
	50					55					60				
Gly	Tyr	Leu	Cys	Ile	Pro	Arg	Thr	Asn	Pro	Val	Tyr	Arg	Gly	Pro	Tyr
65					70					75				80	
Ser	Asn	Pro	Tyr	Ser	Thr	Pro	Tyr	Ser	Gly	Pro	Tyr	Pro	Ala	Ala	Ala
				85					90				95		
Pro	Pro	Leu	Ser	Ala	Pro	Asn	Tyr	Pro	Thr	Ile	Ser	Arg	Pro	Leu	Ile
			100					105					110		
Cys	Arg	Phe	Gly	Tyr	Gln	Met	Asp	Glu	Ser	Asn	Gln	Cys	Val	Asp	Val
		115					120					125			
Asp	Glu	Cys	Ala	Thr	Asp	Ser	His	Gln	Cys	Asn	Pro	Thr	Gln	Ile	Cys
	130					135					140				
Ile	Asn	Thr	Glu	Gly	Gly	Tyr	Thr	Cys	Ser	Cys	Thr	Asp	Gly	Tyr	Trp
145					150					155				160	
Leu	Leu	Glu	Gly	Gln	Cys	Leu	Asp	Ile	Asp	Glu	Cys	Arg	Tyr	Gly	Tyr
				165					170				175		
Cys	Gln	Gln	Leu	Cys	Ala	Asn	Val	Pro	Gly	Ser	Tyr	Ser	Cys	Thr	Cys
			180					185					190		
Asn	Pro	Gly	Phe	Thr	Leu	Asn	Glu	Asp	Gly	Arg	Ser	Cys	Gln	Asp	Val
		195					200					205			
Asn	Glu	Cys	Ala	Thr	Glu	Asn	Pro	Cys	Val	Gln	Thr	Cys	Val	Asn	Thr
	210					215					220				
Tyr	Gly	Ser	Phe	Ile	Cys	Arg	Cys	Asp	Pro	Gly	Tyr	Glu	Leu	Glu	Glu
225					230					235				240	
Asp	Gly	Val	His	Cys	Ser	Asp	Met	Asp	Glu	Cys	Ser	Phe	Ser	Glu	Phe
				245					250				255		
Leu	Cys	Gln	His	Glu	Cys	Val	Asn	Gln	Pro	Gly	Thr	Tyr	Phe	Cys	Ser
			260					265					270		
Cys	Pro	Pro	Gly	Tyr	Ile	Leu	Leu	Asp	Asp	Asn	Arg	Ser	Cys	Gln	Asp
			275				280					285			
Ile	Asn	Glu	Cys	Glu	His	Arg	Asn	His	Thr	Cys	Asn	Leu	Gln	Gln	Thr
	290					295					300				
Cys	Tyr	Asn	Leu	Gln	Gly	Gly	Phe	Lys	Cys	Ile	Asp	Pro	Ile	Arg	Cys
305					310					315				320	
Glu	Glu	Pro	Tyr	Leu	Arg	Ile	Ser	Asp	Asn	Arg	Cys	Met	Cys	Pro	Ala
				325					330				335		
Glu	Asn	Pro	Gly	Cys	Arg	Asp	Gln	Pro	Phe	Thr	Ile	Leu	Tyr	Arg	Asp
			340					345					350		
Met	Asp	Val	Val	Ser	Gly	Arg	Ser	Val	Pro	Ala	Asp	Ile	Phe	Gln	Met
		355					360					365			
Gln	Ala	Thr	Thr	Arg	Tyr	Pro	Gly	Ala	Tyr	Tyr	Ile	Phe	Gln	Ile	Lys
						375					380				
Ser	Gly	Asn	Glu	Gly	Arg	Glu	Phe	Tyr	Met	Arg	Gln	Thr	Gly	Pro	Ile
385					390					395				400	
Ser	Ala	Thr	Leu	Val	Met	Thr	Arg	Pro	Ile	Lys	Gly	Pro	Arg	Glu	Ile
				405					410				415		
Gln	Leu	Asp	Leu	Glu	Met	Ile	Thr	Val	Asn	Thr	Val	Ile	Asn	Phe	Arg
			420					425				430			
Gly	Ser	Ser	Val	Ile	Arg	Leu	Arg	Ile	Tyr	Val	Ser	Gln	Tyr	Pro	Phe
		435					440					445			

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2550 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
 (A) LIBRARY: CORNNOT01
 (B) CLONE: 45517

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CCAAGATTGT	TGTGAGGAGT	CTAGCCAGTT	GGTGAGCGCT	GTAATCTGAA	CCAGCTGTGT	60
CCAGACTGAG	GCCCCATTTG	CATTATTTAA	CATACTTAGA	AAATGAAGTG	TTCATTTTAA	120
ACATTCTCTC	TCCAATTGGT	TTAATGCTGA	ATTACTGAAG	AGGGCTAAGC	AAAACCAGGT	180
GCTTGCGCTG	AGGGCTCTGC	AGTGGCTGGG	AGGACCCCGG	CGCTCTCCCC	GTGTCTCTCTC	240
CACGACTCGC	TCGGCCCCCTC	TGGAATAAAA	CACCCGCGAG	CCCCGAGGGC	CCAGAGGAGG	300
CCGACGTGCC	CGAGCTCCTC	CGGGGGTCCC	GCCCGCGAGC	TTTCTTCTCG	CCTTCGCATC	360
TCCTCCTCGC	GCGTCTTGGA	CATGCCAGGA	ATAAAAAGGA	TACTCACTGT	TACCATTCTG	420
GCTCTCTGTC	TTCCAAGCCC	TGGAATGCA	CAGGCACAGT	GCACGAATGG	CTTTGACCTG	480
GATCGCCAGT	CAGGACAGTG	TTTAGATATT	GATGAATGCC	GAACCATCCC	CGAGGCCTGC	540
CGAGGAGACA	TGATGTGTGT	TAACCAAAAT	GGCGGGTATT	TATGCATTCC	CCGGACAAAC	600
CCTGTGTATC	GAGGGCCCTA	CTCGAACCCC	TACTCGACCC	CCTACTCAGG	TCCGTACCCA	660
GCAGCTGCCC	CACCACTCTC	AGCTCCAAAC	TATCCCACGA	TCTCCAGGCC	TCTTATATGC	720
CGCTTTGGAT	ACCAGATGGA	TGAAAGCAAC	CAATGTGTGG	ATGTGGACGA	GTGTGCAACA	780
GATTCCCACC	AGTGCAACCC	CACCCAGATC	TGCATCAATA	CTGAAGGCGG	GTACACCTGC	840
TCCTGCACCG	ACGGATATTG	GCTTCTGGAA	GGCCAGTGC	TAGACATTGA	TGAATGTCGC	900
TATGGTTACT	GCCAGCAGCT	CTGTGCGAAT	GTTCCTGGAT	CCTATTCTTG	TACATGCAAC	960
CCTGGTTTAA	CCCTCAATGA	GGATGGAAGG	TCTTGCCAAG	ATGTGAACGA	GTGTGCCACC	1020
GAGAACCCCT	GCGTGCAAAC	CTGCGTCAAC	ACCTACGGCT	CTTTCATCTG	CCGCTGTGAC	1080
CCAGGATATG	AACTTGAGGA	AGATGGCGTT	CATTGCAGTG	ATATGGACGA	GTGCAGCTTC	1140
TCTGAGTTCC	TCTGCCAACA	TGAGTGTGTG	AACCAGCCCC	GCACATACTT	CTGCTCCTGC	1200
CCTCCAGGCT	ACATCCTGCT	GGATGACAA	CGAAGCTGCC	AAGACATCAA	CGAATGTGAG	1260
CACAGGAACC	ACACGTGCAA	CTGCGAGCAG	ACGTGCTACA	ATTTACAAGG	GGGCTTCAAA	1320
TGCATCGACC	CCATCCGCTG	TGAGGAGCCT	TATCTGAGGA	TCAGTGATAA	CCGCTGTATG	1380
TGTCCTGCTG	AGAACCCTGG	CTGCAGAGAC	CAGCCCTTTA	CCATCTTGTA	CCGGGACATG	1440
GACGTGGTGT	CAGGACGCTC	CGTTCCCGCT	GACATCTTCC	AAATGCAAGC	CACGACCCGC	1500
TACCCCTGGG	CCTATTACAT	TTTCCAGATC	AAATCTGGGA	ATGAGGGCAG	AGAATTTTAC	1560
ATGCGGCAAA	CGGGCCCCAT	CAGTGCCACC	CTGGTGATGA	CACGCCCCAT	CAAAGGGCCC	1620
CGGGAAATCC	AGCTGGACTT	GGAAATGATC	ACTGTCAACA	CTGTCATCAA	CTTCAGAGGC	1680
AGCTCCGTGA	TCCGACTGCG	GATATATGTG	TCGCAGTACC	CATTCTGAGC	CTCGGGCTGG	1740
AGCCTCCGAC	GCTGCCTCTC	ATTGGCACCA	AGGGACAGGA	GAAGAGAGGA	AATAACAGAG	1800
AGAATGAGAG	CGACACAGAC	GTTAGGCATT	TCCCTGCTGAA	CGTTTCCCCG	AAGAGTCAGC	1860
CCCGACTTCC	TGACTCTCAC	CTGTACTATT	GCAGACCTGT	CACCCTGCAG	GACTTGCCAC	1920
CCCCAGTTCC	TATGACACAG	TTATCAAAAA	GTATTATCAT	TGCTCCCCTG	ATAGAAGATT	1980
GTTGGTGAAT	TTTCAAGGCC	TTCAGTTTAT	TTCCACTATT	TTCAAAGAAA	ATAGATTAGG	2040
TTTGCGGGGG	TCTGAGTCTA	TGTTCAAAAG	CTGTGAACAG	CTTGCTGTCA	CTTCTTCACC	2100
TCTTCCACTC	CTTCTCTCAC	TGTGTTACTG	CTTTGCAAAG	ACCCGGGAGC	TGGCGGGGAA	2160
CCCTGGGAGT	AGCTAGTTTG	CTTTTTGCGT	ACACAGAGAA	GGCTATGTAA	ACAAACCACA	2220
GCAGGATCGA	AGGGTTTTTA	GAGAATGTGT	TTCAAAACCA	TGCCTGGTAT	TTTCAACCAT	2280
AAAAGAAGTT	TCAGTTGTCC	TTAAATTTGT	ATAACGGTTT	AATTCTGTCT	TGTTCAATTT	2340
GAGTATTTTT	AAAAAATATG	TCGTAGAATT	CCTTCGAAAG	GCCTTCAGAC	ACATGCTATG	2400
TTCTGTCTTC	CCAAACCCAG	TCTCCTCTCC	ATTTTAGCCC	AGTGTTTTCT	TTGAGGACCC	2460
CTTAATCTTG	CTTTCTTTAG	AATTTTACC	CAATTGGATT	GGAATGCAGA	GGTCTCCAAA	2520
CTGATTAAAT	ATTTGAAGAG	AAAAAAAAAA				2550

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 540 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
 (A) LIBRARY: BRAITUT13
 (B) CLONE: 1621777

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Gly	Thr	Thr	Ala	Arg	Ala	Ala	Leu	Val	Leu	Thr	Tyr	Leu	Ala	Val
1				5				10						15	
Ala	Ser	Ala	Ala	Ser	Glu	Gly	Gly	Phe	Thr	Ala	Thr	Gly	Gln	Arg	Gln
			20					25					30		
Leu	Arg	Pro	Glu	His	Phe	Gln	Glu	Val	Gly	Tyr	Ala	Ala	Pro	Pro	Ser
		35					40					45			
Pro	Pro	Leu	Ser	Arg	Ser	Leu	Pro	Met	Asp	His	Pro	Asp	Ser	Ser	Gln
		50				55					60				
His	Gly	Pro	Pro	Phe	Glu	Gly	Gln	Ser	Gln	Val	Gln	Pro	Pro	Pro	Ser
65					70					75					80
Gln	Glu	Ala	Thr	Pro	Leu	Gln	Gln	Glu	Lys	Leu	Leu	Pro	Ala	Gln	Leu
				85					90					95	
Pro	Ala	Glu	Lys	Glu	Val	Gly	Pro	Pro	Leu	Pro	Gln	Glu	Ala	Val	Pro
			100					105					110		
Leu	Gln	Lys	Glu	Leu	Pro	Ser	Leu	Gln	His	Pro	Asn	Glu	Gln	Lys	Glu
		115					120					125			
Gly	Met	Pro	Ala	Pro	Phe	Gly	Asp	Gln	Ser	His	Pro	Glu	Pro	Glu	Ser
	130					135					140				
Trp	Asn	Ala	Ala	Gln	His	Cys	Gln	Gln	Asp	Arg	Ser	Gln	Gly	Gly	Trp
145					150					155					160
Gly	His	Arg	Leu	Asp	Gly	Phe	Pro	Pro	Gly	Arg	Pro	Ser	Pro	Asp	Asn
				165					170					175	
Leu	Asn	Gln	Ile	Cys	Leu	Pro	Asn	Arg	Gln	His	Val	Val	Tyr	Gly	Pro
			180					185					190		
Trp	Asn	Leu	Pro	Gln	Ser	Ser	Tyr	Ser	His	Leu	Thr	Arg	Gln	Gly	Glu
		195					200					205			
Thr	Leu	Asn	Phe	Leu	Glu	Ile	Gly	Tyr	Ser	Arg	Cys	Cys	His	Cys	Arg
		210				215					220				
Ser	His	Thr	Asn	Arg	Leu	Glu	Cys	Ala	Lys	Leu	Val	Trp	Glu	Glu	Ala
225					230					235					240
Met	Ser	Arg	Phe	Cys	Glu	Ala	Glu	Phe	Ser	Val	Lys	Thr	Arg	Pro	His
				245					250					255	
Trp	Cys	Cys	Thr	Arg	Gln	Gly	Glu	Ala	Arg	Phe	Ser	Cys	Phe	Gln	Glu
			260					265					270		
Glu	Ala	Pro	Gln	Pro	His	Tyr	Gln	Leu	Arg	Ala	Cys	Pro	Ser	His	Gln
		275					280					285			
Pro	Asp	Ile	Ser	Ser	Gly	Leu	Glu	Leu	Pro	Phe	Pro	Pro	Gly	Val	Pro
		290				295					300				
Thr	Leu	Asp	Asn	Ile	Lys	Asn	Ile	Cys	His	Leu	Arg	Arg	Phe	Arg	Ser
305					310					315					320
Val	Pro	Arg	Asn	Leu	Pro	Ala	Thr	Asp	Pro	Leu	Gln	Arg	Glu	Leu	Leu
				325					330					335	
Ala	Leu	Ile	Gln	Leu	Glu	Arg	Glu	Phe	Gln	Arg	Cys	Cys	Arg	Gln	Gly
			340					345					350		
Asn	Asn	His	Thr	Cys	Thr	Trp	Lys	Ala	Trp	Glu	Asp	Thr	Leu	Asp	Lys
		355					360					365			
Tyr	Cys	Asp	Arg	Glu	Tyr	Ala	Val	Lys	Thr	His	His	His	Leu	Cys	Cys
	370					375					380				
Arg	His	Pro	Pro	Ser	Pro	Thr	Arg	Asp	Glu	Cys	Phe	Ala	Arg	Arg	Ala
385					390					395					400
Pro	Tyr	Pro	Asn	Tyr	Asp	Arg	Asp	Ile	Leu	Thr	Ile	Asp	Ile	Gly	Arg
			405						410					415	
Val	Thr	Pro	Asn	Leu	Met	Gly	His	Leu	Cys	Gly	Asn	Gln	Arg	Val	Leu
			420					425					430		

```

Thr Lys His Lys His Ile Pro Gly Leu Ile His Asn Met Thr Ala Arg
      435                      440                      445
Cys Cys Asp Leu Pro Phe Pro Glu Gln Ala Cys Cys Ala Glu Glu Glu
      450                      455                      460
Lys Leu Thr Phe Ile Asn Asp Leu Cys Gly Pro Arg Arg Asn Ile Trp
      465                      470                      475                      480
Arg Asp Pro Ala Leu Cys Cys Tyr Leu Ser Pro Gly Asp Glu Gln Val
      485                      490                      495
Asn Cys Phe Asn Ile Asn Tyr Leu Arg Asn Val Ala Leu Val Ser Gly
      500                      505                      510
Asp Thr Glu Asn Ala Lys Gly Gln Gly Glu Gln Gly Ser Thr Gly Gly
      515                      520                      525
Thr Asn Ile Ser Ser Thr Ser Glu Pro Lys Glu Glu
      530                      535                      540

```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1899 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: BRAITUT13
 (B) CLONE: 162177

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

TGGGTGCAAG CTCACAACCG TAACAGCCAC CAGACAAGCT TCAGTGGCCG GCCCTTCACA      60
TCCAGACTTG CCTGAGAGGA CCCACCTCTG AGTGTCAGT GGTCAAGTTGC CCCAGGATGG      120
GGACCACAGC CAGAGCAGCC TTGGTCTTGA CCTATTTGGC TGTTGCTTCT GCTGCCCTCTG      180
AGGGAGGCTT CACGGCTACA GGACAGAGGC AGCTGAGGCC AGAGCACTTT CAAGAAGTTG      240
GCTACGCAGC TCCCCCTCC CCACCCCTAT CCCGAAGCCT CCCCATGGAT CACCCTGACT      300
CCTCTCAGCA TGGCCCTCCC TTTGAGGGAC AGAGTCAAGT GCAGCCCCCT CCCTCTCAGG      360
AGGCCACCCC TCTCCAACAG GAAAAGCTGC TACCTGCCCA ACTCCCTGCT GAAAAGGAAG      420
TGGGTCCCCC TCTCCCTCAG GAAGCTGTCC CCTTCCAAA AGAGCTGCCC TCTCTCCAGC      480
ACCCCAATGA ACAGAAGGAA GGAATGCCAG ATCCATTGGG GGACCAGAGC CATCCAGAAC      540
CTGAGTCTTG GAATGCAGCC CAGCACTGCC AACAGGACCG GTCCCAAGGG GGCTGGGGCC      600
ACCGGCTGGA TGGCTTCCCC CTTGGGCGGC CTTCTCCAGA CAATCTGAAC CAAATCTGCC      660
TTCCTAACCG TCAGCATGTG GTATATGGTC CCTGGAACCT ACCACAGTCC AGCTACTCCC      720
ACCTCACTCG CCAGGGTGAG ACCCTCAATT TCCTGGAGAT TGGATATTCC CGCTGCTGCC      780
ACTGCCGCAG CCACACAAAC CGCCTAGAGT GTGCCAAACT TGTGTGGGAG GAAGCAATGA      840
GCCGATTCTG TGAGGCCGAG TTCTCGGTCA AGACCCGACC CCACTGGTGC TGCACGCGGC      900
AGGGGAGGC TCGTTCTCC TGCTTCCAGG AGGAAGCTCC CCAGCCACAC TACCAGCTCC      960
GGGCCTGCCC CAGCCATCAG CCTGATATTT CCTCGGGTCT TGAGCTGCCT TTCCCTCCTG      1020
GGGTGCCAC ATTGGACAAT ATCAAGAACA TCTGCCACCT GAGGCGCTTC CGCTCTGTGC      1080
CACGCAACCT GCCAGCTACT GACCCCTAC AAAGGGAGCT GCTGGCACTG ATCCAGCTGG      1140
AGAGGGAGTT CCAGCGCTGC TGCCGCCAGG GGAACAATCA CACCTGTACA TGGAAGGCCT      1200
GGGAGGATAC CCTTGACAAA TACTGTGACC GGGAGTATGC TGTGAAGACC CACCACCACT      1260
TGTGTTGCCG CCACCCTCCC AGCCCTACTC GGGATGAGTG GAGACACTGA CTTTGCCCGT      1320
ACCCCAACTA TGACCGGGAC ATCTTGACCA TTGACATCGG TCGAGTCACC CCAACCTCA      1380
TGGGCCACCT CTGTGGAAAC CAAAGAGTTC TCACCAAGCA TAAACATATT CCTGGGCTGA      1440
TCCACAACAT GACTGCCC GC TGCTGTGACC TGCCATTTCC AGAACAGGCC TGCTGTGCAG      1500
AGGAGGAGAA ATTAACCTTC ATCAATGATC TGTGTGGTCC CCGACGTAAC ATCTGGCGAG      1560
ACCCTGCCCC CTGCTGTTAC CTGAGTCTCT GGGATGAACA GGTCAACTGC TTCAACATCA      1620
ATTATCTGAG GAACGTGGCT CTAGTGTCTG GAGACACTGA GAACGCCAAG GGCCAGGGG      1680
AGCAGGGCTC AACTGGAGGA ACAAATATCA GCTCCACCTC TGAGCCCAAG GAAGAATGAG      1740
TCACCCAGA GCCCTAGAGG GTCAGATGGG GGAACCCCA CCCTGCCCCA CCCATCTGAA      1800
CACTCATTAC ACTAAACACC TCTTGGATTT GGTGTCCTCA TTGTCTATCT AATGTCTCAC      1860
CCGCAGTGTT TTAAGTGGAT CTTGGTGCCC TGGCCCAGG      1899

```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 387 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
 (B) CLONE: 458228

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Ala	Thr	Ser	Gly	Val	Leu	Pro	Gly	Gly	Gly	Phe	Val	Ala	Ser	Ala	1	5	10	15
Ala	Ala	Val	Ala	Gly	Pro	Glu	Met	Gln	Thr	Gly	Arg	Asn	Asn	Phe	Val	20	25	30	
Ile	Arg	Arg	Asn	Pro	Ala	Asp	Pro	Gln	Arg	Ile	Pro	Ser	Asn	Pro	Ser	35	40	45	
His	Arg	Ile	Gln	Cys	Ala	Ala	Gly	Tyr	Glu	Gln	Ser	Glu	His	Asn	Val	50	55	60	
Cys	Gln	Asp	Ile	Asp	Glu	Cys	Thr	Ala	Gly	Thr	His	Asn	Cys	Arg	Ala	65	70	75	80
Asp	Gln	Val	Cys	Ile	Asn	Leu	Arg	Gly	Ser	Phe	Ala	Cys	Gln	Cys	Pro	85	90	95	
Pro	Gly	Tyr	Gln	Lys	Arg	Gly	Glu	Gln	Cys	Val	Asp	Ile	Asp	Glu	Cys	100	105	110	
Thr	Ile	Pro	Pro	Tyr	Cys	His	Gln	Arg	Cys	Val	Asn	Thr	Pro	Gly	Ser	115	120	125	
Phe	Tyr	Cys	Gln	Cys	Ser	Pro	Gly	Phe	Gln	Leu	Ala	Ala	Asn	Asn	Tyr	130	135	140	
Thr	Cys	Val	Asp	Ile	Asn	Glu	Cys	Asp	Ala	Ser	Asn	Gln	Cys	Ala	Gln	145	150	155	160
Gln	Cys	Tyr	Asn	Ile	Leu	Gly	Ser	Phe	Ile	Cys	Gln	Cys	Asn	Gln	Gly	165	170	175	
Tyr	Glu	Leu	Ser	Asp	Arg	Leu	Asn	Cys	Glu	Asp	Ile	Asp	Glu	Cys		180	185	190	
Arg	Thr	Ser	Ser	Tyr	Leu	Cys	Gln	Tyr	Gln	Cys	Val	Asn	Glu	Pro	Gly	195	200	205	
Lys	Phe	Ser	Cys	Met	Cys	Pro	Gln	Gly	Tyr	Gln	Val	Val	Arg	Ser	Arg	210	215	220	
Thr	Cys	Gln	Asp	Ile	Asn	Glu	Cys	Glu	Thr	Thr	Asn	Glu	Cys	Arg	Glu	225	230	235	240
Asp	Glu	Met	Cys	Trp	Asn	Tyr	His	Gly	Gly	Phe	Arg	Cys	Tyr	Pro	Arg	245	250	255	
Asn	Pro	Cys	Gln	Asp	Pro	Tyr	Ile	Leu	Thr	Pro	Glu	Asn	Arg	Cys	Val	260	265	270	
Cys	Pro	Val	Ser	Asn	Ala	Met	Cys	Arg	Glu	Leu	Pro	Gln	Ser	Ile	Val	275	280	285	
Tyr	Lys	Tyr	Met	Ser	Ile	Arg	Ser	Asp	Arg	Ser	Val	Pro	Ser	Asp	Ile	290	295	300	
Phe	Gln	Ile	Gln	Ala	Thr	Thr	Ile	Tyr	Ala	Asn	Thr	Ile	Asn	Thr	Phe	305	310	315	320
Arg	Ile	Lys	Ser	Gly	Asn	Glu	Asn	Gly	Glu	Phe	Tyr	Leu	Arg	Gln	Thr	325	330	335	
Ser	Pro	Val	Ser	Ala	Met	Leu	Val	Leu	Val	Lys	Ser	Leu	Ser	Gly	Pro	340	345	350	
Arg	Glu	His	Ile	Val	Asp	Leu	Glu	Met	Leu	Thr	Val	Ser	Ser	Ile	Gly	355	360	365	

Thr Phe Arg Thr Ser Ser Val Leu Arg Leu Thr Ile Ile Val Gly Pro
 370 375 380
 Phe Ser Phe
 385

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 559 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
 (B) CLONE: 496120

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Gly Thr Val Ser Arg Ala Ala Leu Ile Leu Ala Cys Leu Ala Leu
 1 5 10 15
 Ala Ser Ala Ala Ser Glu Gly Ala Phe Lys Ala Ser Asp Gln Arg Glu
 20 25 30
 Met Thr Pro Glu Arg Leu Phe Gln His Leu His Glu Val Gly Tyr Ala
 35 40 45
 Ala Pro Pro Ser Leu Pro Gln Thr Arg Arg Leu Arg Val Asp His Ser
 50 55 60
 Val Thr Ser Leu His Asp Pro Pro Leu Phe Glu Glu Gln Arg Glu Val
 65 70 75 80
 Gln Pro Pro Ser Ser Pro Glu Asp Ile Pro Val Tyr Glu Glu Asp Trp
 85 90 95
 Pro Thr Phe Leu Asn Pro Asn Val Asp Lys Ala Gly Pro Ala Val Pro
 100 105 110
 Gln Glu Ala Ile Pro Leu Gln Lys Glu Gln Pro Pro Pro Gln Val His
 115 120 125
 Ile Glu Gln Lys Glu Ile Asp Pro Pro Ala Gln Pro Gln Glu Glu Ile
 130 135 140
 Val Gln Lys Glu Val Lys Pro His Thr Leu Ala Gly Gln Leu Pro Pro
 145 150 155 160
 Glu Pro Arg Thr Trp Asn Pro Ala Arg His Cys Gln Gln Gly Arg Arg
 165 170 175
 Gly Val Trp Gly His Arg Leu Asp Gly Phe Pro Pro Gly Arg Pro Ser
 180 185 190
 Pro Asp Asn Leu Lys Gln Ile Cys Leu Pro Glu Arg Gln His Val Ile
 195 200 205
 Tyr Gly Pro Trp Asn Leu Pro Gln Thr Gly Tyr Ser His Leu Ser Arg
 210 215 220
 Gln Gly Glu Thr Leu Asn Val Leu Glu Thr Gly Tyr Ser Arg Cys Cys
 225 230 235 240
 Pro Cys Arg Ser Asp Thr Asn Arg Leu Asp Cys Leu Lys Leu Val Trp
 245 250 255
 Glu Asp Ala Met Thr Gln Phe Cys Glu Ala Glu Phe Ser Val Lys Thr
 260 265 270
 Arg Pro His Leu Cys Cys Arg Leu Arg Gly Glu Glu Arg Phe Ser Cys
 275 280 285
 Phe Gln Lys Glu Ala Pro Arg Pro Asp Tyr Leu Leu Arg Pro Cys Pro
 290 295 300
 Val His Gln Asn Gly Met Ser Ser Gly Pro Gln Leu Pro Phe Pro Pro
 305 310 315 320
 Gly Leu Pro Thr Pro Asp Asn Val Lys Asn Ile Cys Leu Leu Arg Arg
 325 330 335

1978/79	1979/80	1980/81	1981/82	1982/83	1983/84	1984/85	1985/86	1986/87	1987/88	1988/89	1989/90	1990/91	1991/92	1992/93	1993/94	1994/95	1995/96	1996/97	1997/98	1998/99	1999/00	2000/01	2001/02	2002/03	2003/04	2004/05	2005/06	2006/07	2007/08	2008/09	2009/10	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16	2016/17	2017/18	2018/19	2019/20	2020/21	2021/22	2022/23	2023/24	2024/25	2025/26	2026/27	2027/28	2028/29	2029/30	2030/31	2031/32	2032/33	2033/34	2034/35	2035/36	2036/37	2037/38	2038/39	2039/40	2040/41	2041/42	2042/43	2043/44	2044/45	2045/46	2046/47	2047/48	2048/49	2049/50	2050/51	2051/52	2052/53	2053/54	2054/55	2055/56	2056/57	2057/58	2058/59	2059/60	2060/61	2061/62	2062/63	2063/64	2064/65	2065/66	2066/67	2067/68	2068/69	2069/70	2070/71	2071/72	2072/73	2073/74	2074/75	2075/76	2076/77	2077/78	2078/79	2079/80	2080/81	2081/82	2082/83	2083/84	2084/85	2085/86	2086/87	2087/88	2088/89	2089/90	2090/91	2091/92	2092/93	2093/94	2094/95	2095/96	2096/97	2097/98	2098/99	2099/00	2100/01	2101/02	2102/03	2103/04	2104/05	2105/06	2106/07	2107/08	2108/09	2109/10	2110/11	2111/12	2112/13	2113/14	2114/15	2115/16	2116/17	2117/18	2118/19	2119/20	2120/21	2121/22	2122/23	2123/24	2124/25	2125/26	2126/27	2127/28	2128/29	2129/30	2130/31	2131/32	2132/33	2133/34	2134/35	2135/36	2136/37	2137/38	2138/39	2139/40	2140/41	2141/42	2142/43	2143/44	2144/45	2145/46	2146/47	2147/48	2148/49	2149/50	2150/51	2151/52	2152/53	2153/54	2154/55	2155/56	2156/57	2157/58	2158/59	2159/60	2160/61	2161/62	2162/63	2163/64	2164/65	2165/66	2166/67	2167/68	2168/69	2169/70	2170/71	2171/72	2172/73	2173/74	2174/75	2175/76	2176/77	2177/78	2178/79	2179/80	2180/81	2181/82	2182/83	2183/84	2184/85	2185/86	2186/87	2187/88	2188/89	2189/90	2190/91	2191/92	2192/93	2193/94	2194/95	2195/96	2196/97	2197/98	2198/99	2199/00	2200/01	2201/02	2202/03	2203/04	2204/05	2205/06	2206/07	2207/08	2208/09	2209/10	2210/11	2211/12	2212/13	2213/14	2214/15	2215/16	2216/17	2217/18	2218/19	2219/20	2220/21	2221/22	2222/23	2223/24	2224/25	2225/26	2226/27	2227/28	2228/29	2229/30	2230/31	2231/32	2232/33	2233/34	2234/35	2235/36	2236/37	2237/38	2238/39	2239/40	2240/41	2241/42	2242/43	2243/44	2244/45	2245/46	2246/47	2247/48	2248/49	2249/50	2250/51	2251/52	2252/53	2253/54	2254/55	2255/56	2256/57	2257/58	2258/59	2259/60	2260/61	2261/62	2262/63	2263/64	2264/65	2265/66	2266/67	2267/68	2268/69	2269/70	2270/71	2271/72	2272/73	2273/74	2274/75	2275/76	2276/77	2277/78	2278/79	2279/80	2280/81	2281/82	2282/83	2283/84	2284/85	2285/86	2286/87	2287/88	2288/89	2289/90	2290/91	2291/92	229
---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	-----

111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929